UNDERSTANDING SPATIAL DATA USABILITY

Gary J. Hunter[1], Monica Wachowicz[2] and Arnold K. Bregt[2]

*[*1] Department of Geomatics, University of Melbourne, Victoria 3010, Australia
Email: garyh@unimelb.edu.au*

*[2] Centre for Geo-Information, Wageningen UR, PO Box 339, 6700 AH Wageningen, The Netherlands
Email: a.k.bregt@alterra.wag-ur.nl, m.wachowicz@alterra.wag-ur.nl*

ABSTRACT

In recent geographical information science literature, a number of researchers have made passing reference to an apparently new characteristic of spatial data known as 'usability'. While this attribute is well-known to professionals engaged in software engineering and computer interface design and testing, extension of the concept to embrace information would seem to be a new development. Furthermore, while notions such as the use and value of spatial information, and the diffusion of spatial information systems, have been the subject of research since the late-1980s, the current references to usability clearly represent something which extends well beyond that initial research. Accordingly, the purposes of this paper are: (1) to understand what is meant by spatial data usability; (2) to identify the elements that might comprise usability; and (3) to consider what the related research questions might be.

Keywords: Spatial data, Usability, Usability elements, Research questions.

1 INTRODUCTION

> *The information you have is not the information you want.*
> *The information you want is not the information you need.*
> *The information you need is not the information you can obtain.*
> *The information you can obtain costs more than you want to pay.*

(anon., as reported in Bernstein, 1998, p. 202)

It is now almost 150 years since Dr John Snow combined spatial data showing the locations of cholera deaths in London with those of water pumps, thereby helping to prove his theory about the source and transmission of an outbreak of the deadly disease that claimed 600 lives in its first ten days (UCLA, 2001). That famous example is now taught to students worldwide in fields such as geography, public health and epidemiology, and serves as a perfect example of how spatial data can be very effectively applied in critical situations.

Moving to the present day, there are other applications of spatial data which, although they do not have the same life-saving impact as Snow's work, are nevertheless proving to be extremely valuable to the communities in which they are applied. For instance, the internationally-recognised South Australian Land Ownership and Tenure System (LOTS) has grown from handling a few hundred inquiries per week in the late-1980s, to now receiving over 2.5 million fee-based, on-line queries per annum for property information from more than 1500 public and private terminals distributed throughout a State that has fewer than 800,000 land parcels (Government of South Australia, 2001).

On the other hand, a bold multi-million dollar initiative of the 1970s to provide on-line interactive color maps of statistical data as part of the U.S. White House information system (the Domestic Information Display System—DIDS), ended in complete abandonment within the space of a few short years (Cowen, 1982). Similarly, efforts by census officials in Israel to make the 1995 population census data (at the rare, individual household-level) both publicly and freely available to interest-groups such as academics, planners and marketing agencies, have met with an almost complete lack of user response (Benenson, I., pers. comm., November 2001).

Elsewhere, in the area of environmental impact assessment, there have been cases reported in the past few years where government officials abandoned major projects essentially because they were unwilling to proceed with their decision-making due to overwhelming concerns about the validity of the spatial information presented to them (Beven, 2000). In one instance relating to the proposed establishment of a deep repository for radioactive waste material at Sellafield in the UK—the former Windscale site—"Simulations of the possible groundwater flows differed drastically between modelers on both sides of the argument" (Beven, 2000, p. 2605), while at the same time there have been similar cases in the U.S. "…where the ability of hydrological models to make useful predictions has been called into question" (Beven, 2000, p. 2605).

So there would seem to be a common connection between these examples, involving some fundamental characteristic that has resulted in these spatial data applications being either very successful or unsuccessful. As such, it would appear they all demonstrate very high or very low degrees of data 'usefulness' or 'usability', which in turn has produced very positive or negative economic, social, political, environmental or scientific impacts.

With this in mind, our motivation here lies in questioning exactly what it is that distinguishes these cases from other, more mundane, examples. For instance, is low usability caused solely by the wrong combination of data, algorithms or models for a given application, or is it simply a matter of poor data quality? Alternatively, is high usability proportional to the degree of 'interestingness' in the data (as data miners or knowledge discoverers would say), or perhaps the by-product of data integration and value-adding processes? Or could it be these differences are caused simply by some unpredictable, indescribable phenomenon that produces such extreme examples?

At this stage we do not know, but given the very large expenditure of human and capital resources nowadays on the development of spatial data products, both for public good and commercial purposes, it would seem to be a goal worth pursuing to ensure they are as made as 'usable' as possible. Clearly, with a better understanding of usability we might be able to increase the number of 'successes' and reduce the incidence of 'failures' in the development and application of spatial datasets.

Accordingly, to help gain a better understanding of spatial data usability this paper is arranged as follows. Section 2 examines what is meant by usability by looking at its origins and subsequent advances in the wider field of information science and technology. Section 3 provides a preliminary assessment of what the elements might be that comprise usability in the context of spatial data. Finally, Section 4 considers what the relevant research questions and priorities might be in this field, and Section 5 provides some concluding remarks.

2 INTRODUCING USABILITY

Referring back to the quotation at the beginning of this paper, as supplied to Bernstein (1998) by an anonymous colleague, the somewhat cynical view being taken there is that perfect information can never be obtained, and the reason it is unachievable is because we are inevitably forced to make compromises between the means we have at our disposal to perform a task and what we would like to have. Of course few things in life are perfect, but nonetheless we in fact do encounter situations where we believe we have sufficient information needed to satisfactorily complete the tasks expected of us. So could we perhaps put a more positive 'spin' on our anonymous sage's comments and, accepting that perfect information is a rarity, ask whether there are certain fundamental elements that need to be present in information for it to be considered sufficiently usable for our purposes?

2.1 Usability and information technology

To help answer this question we need to go back and examine the interaction between humans and technology—and in particular, information technology. Research into how we use this technology is relatively new, essentially due to the tremendous pace of developments that have taken place over the last 50 years. Of course, hardware and software usability was not a key issue for the highly skilled scientists who developed the world's first computers in the 1940s and 50s, since they were the same people who would use this new technology—and they knew they would inevitably encounter problems that would have to be overcome simply as part of the process of scientific discovery. However, the use of information technology is no longer restricted to secretive research laboratories, and instead has become widely available to the broader community which has in turn brought the need to ensure it is made as useful and usable as possible.

Since those early experimental days of computer science, information technology has grown to become an integral part of our daily lives in fields as diverse as business, finance, telecommunication, transport, health, defense and education. However, in return for our acceptance of its widespread impact (either willingly or unwillingly), we have come to expect (and increasingly demand) that information technology will perform the functions for which it is designed in an acceptable manner. Unfortunately, these new technologies can still have their teething troubles, and examples which many of us would be familiar include:

- frustrating software 'crashes';
- losing a bank card to an automatic teller machine;
- problems encountered in programming a video recorder;
- difficulty in remembering how to access the many functions on mobile telephones;
- the 'phone rage' that can be experienced with automated voice response services; and
- being unable to solve a problem when using a software Help 'wizard'.

Of course, it was never intended that these problems should occur and product developers naturally aim to ensure their devices are as easy to use as possible by consumers. From the developer's perspective, customers who are frustrated with products that are difficult to use or produce unexpected/erroneous results may (1) return the item and seek a refund on its purchase, (2) tell their colleagues about their disappointment and dissatisfaction, and (3) are unlikely to be repeat clients. This can all directly impact upon a company's success—after all, no-one wants to be known in business as having designed a 'lemon', and we regularly see examples in the media of poor usability when a product is recalled for replacement or repair, or worse, removed entirely from the marketplace. So the incentive to 'get things right' in a competitive, commercial environment is very real—and this even includes the field of Geographic Information Systems (or GIS) (for example, see ESRI, 2001)

Perhaps the most recent and highly-publicized case of poor technological usability concerned the U.S. Presidential election held in 2000, in which the ballot card design and punch-card technology used in some parts of that nation were considered by some people to be responsible for voting irregularities leading to a series of costly and controversial court cases (and even, perhaps, to the wrong candidate being elected President). The usability problem here related first to the way candidates names were arranged in a 2-column (or 'butterfly') format rather than in a single list. Between the two columns of names, a single column of punch-card tabs was aligned—with left and right arrows pointing to the correct tab to punch for each pair of presidential/vice-presidential nominees. Some elderly voters found the design confusing and believe they may have mistakenly voted for the wrong candidate (STC, 2001). Secondly, the technology used to count the voters' punch-cards was thought to suffer from reliability problems when partly-punched tabs were encountered.

Meanwhile in other fields of information technology, many mobile devices are still not yet seen to be as usable as they might be. For instance, in the area of mobile telephones researchers consider that hands-free usage in vehicles is not that much safer than manual operation because of the inherent distraction caused during driving (Alertbox, 2001a). Usability problems are also experienced with e-mail discussions lists, as observed by the common plea of "How do I get off this list?". In this instance, the problem is often caused by the fact that the instructions for leaving the list came only with the first message acknowledging list subscription—which may have been deleted several years before. Another example of poor usability that can occur when subscribing to e-mail lists is when users are required to untick a box to indicate they do not want to receive promotional messages. Usability testing has shown it is much better to leave these boxes unchecked, so that users must take the conscious action of indicating they want to receive further material, which ensures they do not receive unwanted advertising which in turn can cause them to think the list moderator is abusing their subscription (Alertbox, 2000).

Sometimes usability problems make good comedy, like the famous line in the popular 1970s and 80s U.S. television series "M.A.S.H.", where the military instructions for defusing a bomb went something along the lines of "Begin by cutting the red wire, followed by the blue wire … but FIRST cut the green wire". Of course we can all see the humorous side of such a story, but in everyday life the impact of usability problems can be deadly serious—such as when failures happen in nuclear power plants or air traffic control towers. There are even usability problems being observed in new entertainment technologies such as virtual reality equipment, where undesirable side effects like motion sickness can occur after only a few minutes use.

While there are obvious downsides to poor technological usability, there are equally beneficial aspects to it as well, and an obvious example of this occurred with the development of the Macintosh computer some 20 years ago. Until its invention, computer operating systems had been command-line based and their application was confined to a relatively small group of specialised and dedicated users. However the introduction of the

Macintosh's graphic user interface with its revolutionary use of icons and 'click and drag' functionality helped open the way for computers to be introduced to mass-markets. In other words, this new form of human-computer interaction resulted in a much more user-friendly product evolving that had immediate appeal to a wider group of consumers.

So the notion of usability is something we are quite familiar with, but what exactly do we mean by it? Typical dictionary definitions show 'usability' to be a noun to the adjective 'usable', which means that something is capable of being used, or is convenient and practicable for use. In the case of information technology, Jacobsen (1999) suggests if we consider any technological system in a general sense, then it would seem reasonable for its success that it should possess the qualities of both utility and usability. In this case, utility is a measure of how well the system completes the tasks for which it is designed, while Jacobsen goes on to note that usability is officially defined in the ISO 9241-11 standard relating to Visual Display Terminal ergonomics to be a combination of effectiveness, efficiency and satisfaction (International Organisation for Standardisation, 1998). That standard reports:

"System usability comprises the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use where:

- *Effectiveness measures the accuracy and completeness with which users achieve specified goals;*
- *Efficiency measures the resources expended in relation to the accuracy and completeness with which users achieve goals; and*
- *Satisfaction measures the freedom from discomfort, and positive attitudes towards the use of the product."*

In a similar vein, ISO/IEC 9126: Software engineering—Product quality (International Organisation for Standardisation, 2001) categorizes usability as a fundamental characteristic of good software and defines it as being:

"The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions.

- *Understandability: the capability of the software product to enable the user to understand whether the software is suitable, and how it can be used for particular tasks and conditions of use.*
- *Learnability: The capability of the software product to enable the user to learn its application.*
- *Operability: The capability of the software product to enable the user to operate and control it."*

Other meanings of usability appear to be variations on these official definitions, as in "[usability] is the measure of the quality of the user's experience when interacting with a product or system—whether a website, a software application, mobile technology, or any user-operated device" (National Cancer Institute, 2001). In addition, new elements are sometimes introduced to usability descriptions, such as 'memorability' (that is, how easily repeat users of the system can re-apply the system after a period of non-use) and 'error frequency and severity' (how often errors are made by users, how serious they are, and how easily users find it to extract themselves from an error and return to operating the system) (National Cancer Institute, 2001; ESRI, 2001).

The early research into usability was not originally known as such, and many of the world's telephone research laboratories conducted psychological and physiological testing of their equipment as far back as the 1960s. Much of that work led in due course to the development of Human-Computer Interface (HCI) studies which had the goal of maximizing the way users interacted with their systems. Increasingly, major companies started to develop their own sets of human factor guidelines based on past experiences and experiments that observed people using their technology—and companies associated with spatial information like Intergraph are known to have published HCI guidelines and distributed them to clients in the late 1980s.

In the area of information technology, much of the usability literature today focuses on evaluating hardware and software, user interfaces and Internet websites. In the case of websites, determining their usability has become critical to ensuring their success—particularly since studies have shown web users cannot find the information they want approximately 60% of the time, which in turn leads to "wasted time, reduced productivity, increased frustration, and a loss of repeat visits and money" (National Cancer Institute, 2001). For evaluation purposes, some website usability engineers believe between five and eight users are all that is needed to detect approximately 85% of the problems present in using a website (Spool, 2001).

Obviously, for some websites such as those containing general public information, it remains a challenge for website designers to measure usability. However, it is clear that where such websites have “content that is highly desirable and unavailable from any other source”, then users will obviously find the site attractive (Spool, 2001). On the other hand, for websites designed for e-commerce purposes, usability success can be directly measured in terms of sales from the site—and the value of e-commerce is certainly recognised by companies such the major U.S. office supplier, Staples, which generated approximately 42% of its $451 million revenue in 2000 through its website.

To help appraise the usability of websites there is now a wide range of predictors or metrics in use—particularly for retail websites—such as the time required to complete a task, the error rate, and the user satisfaction (measured, for example, by repeat visits). For instance, testing a user’s ability to search and order a book through an on-line supplier, or else ordering a certain quantity and type of flower to be delivered in time for Mother’s Day, are quite specific web-based services that can be tested for usability in hard, quantifiable terms (Alertbox, 2001b). In turn, to assess these metrics there is a wide range of usability evaluation methods that can be applied, including performance testing, beta test sites, expert review, cognitive walk-throughs, structured heuristic evaluations, user satisfaction questionnaires, field observation, focus groups and interviews (Hom, 1996; Jacobsen, 1999). However, it should be noted that exactly how these methods are employed to produce usable metrics is seldom revealed for commercial-in-confidence reasons by the consultants who use them.

Nevertheless, there are numerous texts on the subject and usability testing is now taught to information systems students around the world. As such, it comes as no surprise to find there is an association of usability professionals which has operated in the U.S. for the past 10 years, and which functions in the same way and with similar goals as our own spatial information associations (UPA, 2001). For further reading, an example of a commercial usability evaluation of an on-line banking system is described in detail in System Concepts (2001). The problems detected in that particular case study included very slow response times, lack of adherence to standards, form preferred ahead of content, bad customer support, and the absence of a design that met customer needs.

2.2 Usability and non-spatial information

With the notion of usability having become firmly established in the information technology sector it seems natural that it should be extended to cover data and information. In this regard, agencies such as NASA have become concerned with usability from the standpoint of retaining seminal data for archival purposes. From past experience, NASA has found it is sometimes necessary to re-use old data in new ways or new combinations because of the insight it might provide in previously unanticipated circumstances. As such, the observational data of NASA “...represent an asset which must be retained in a usable state into the indefinite future” (KING, 1998).

Similarly, professional archivists are confronted on a daily basis with the critical problem of ensuring that the media on which the digital records of human endeavor are stored, actually remain capable of being maintained in an appropriate manner for future use. For example, the U.S. National Archives and Record Administration (NARA) has long held an interest in the technical development of archival-quality magnetic and optical data storage devices. Accordingly, NARA considers there are three critical factors that affect the future usability of data—life expectancy, data degradation, and technology migration strategies (NARA, 2001). Each of these usability elements is considered to be of equal importance, since there is little point in archiving data on a device that (1) may fail at an unknown point in the future, (2) slowly degrade the quality of its content through the introduction of data errors caused by insufficient or inappropriate maintenance of the storage device, or else (3) be incapable of later transferring data to new storage technologies.

Interestingly, the U.S. Office of Management and Budget (OMB) has recently introduced guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information to all users, both internal and external to U.S. Federal agencies. While the guidelines are primarily designed to allow affected citizens to seek and obtain correction of personal information maintained and disseminated by such agencies, they are also aimed at ensuring a certain minimum level of usability of scientific and statistical data in terms of integrity and reliability. Without these essential characteristics, it is felt by the guideline creators that since much of this information is now distributed far more widely throughout the community via the Internet than previously, there could be an increase in “...the potential harm that can result from the dissemination of information that does not meet basic information quality guidelines" (OMB, 2001). Obvious examples of this are government-produced weather reports, air travel advisory notices, and public health and security alerts.

2.3 Usability and spatial information

Government agencies are also starting to focus on the usability of spatial data. For instance, the U.S. Environmental Protection Agency (EPA) is concerned with data usability primarily because it wants any decisions it takes concerning environmentally-degraded land to be legally defensible. So from its perspective, data usability directly impacts upon the regulatory process. In such cases, this capability will depend upon factors such as sampling event planning, sample collection and data assessment as to whether it is of the correct type and quality to support a regulatory decision (EPA, 2001). Elsewhere in the U.S., the Federal Geographic Data Committee has commissioned usability evaluations of typical clearinghouse websites that are used as a means of accessing spatial data and "…to establish what the primary barriers to usability are, and to identify any existing approaches that seem promising for improving the usability of [such sites]" (Metadata Humane Society, n.d.).

Similarly, the U.S. National Ocean and Atmospheric Administration (NOAA) is examining how its large image holdings, which are often used by government organizations for resource and emergency response mapping, can be more effectively made available. While data such as Digital Orthophoto Quads (DOQs) are extremely valuable for these types of tasks, NOAA has found there can be problems for users in quickly accessing large images via the Internet. This has led to a joint Federal/State government project in high-performance computing and communications that will see enhanced image compression techniques used to address current usability concerns relating to the ease of searching, browsing and retrieving DOQs from on-line libraries (NOAA, 2001).

Finally, there can even be usability problems with the simple printouts of directions from vehicle navigation systems and route planning software packages. As Frank (2001) notes, the presentation of route directions in tabular form can cause confusion for drivers since it is not always clear exactly when an action such as turning off a road should be taken. For instance, what is the exact meaning of "Gloggnitz—after 4.2 km—turn left"? (Frank (2001, p. 73). He considers that such problems are due to the inherent semantics associated with documenting travel instructions. In other words, the way the driver interprets and reacts to the instructions are not necessarily viewed the same way by the software designers who develop these systems. While these differences are no doubt unintentional, they are nonetheless quite real and need to be overcome.

From these examples the notion of usability is already starting to emerge as something that can affect spatial data, however it remains to be seen if the elements that comprise it can be usefully identified.

3 THE ELEMENTS OF SPATIAL DATA USABILITY

Turning our attention now to the possible elements that comprise spatial data usability, it is logical to ask at this stage whether or not the idea of usability is something that has already been dealt with a decade ago in other research efforts—most notably by the U.S. National Center for Geographic Information and Analysis (NCGIA) and the European GISDATA foundation. Researchers involved in those initiatives worked on the topics of 'use and value' and 'diffusion of GIS', and as such examined issues relating to the impediments to GIS adoption, the cost of information, models for assessing the value of information, willingness to pay for information, and the effects of improved information on decision making (NCGIA 1989a; 1989b; 1993). Clearly, many of those topics will have a bearing on data usability, but as will be observed from the rest of this section we believe usability extends beyond the organizational structures and information economics necessary to promote the use of GIS and associated data. Instead, the use of GIS and the development of datasets has grown rapidly throughout the world, so while there is no doubt as to their diffusion there continues to be the lingering issue of ensuring their success.

To begin with, an obvious usability factor would seem to be the **type of application** for which the data are used. Of course, this does not mean that only highly profitable, 'killer' commercial applications can yield high usability, a successful non-commercial example of this are the internet-based weather alert systems that provide a critical public service to vulnerable communities in the U.S. (Earthwatch, 2001). Furthermore, in some cases it is the **adding of value** to spatial information through data **integration** that increases the usability of an application. An instance of this is the now-common linkage of electronic telephone directories to web-based street mapping products, which are causing communities to replace their hardcopy telephone directories with digital alternatives that offer an extended range of information services such as local transport timetables, trip-planning and entertainment guides (YellowMap AG, 2001).

Another example of spatial data integration being responsible for enhanced usability goes back almost two decades to the mid-1980s, when the internationally-recognised South Australian Land Ownership and Tenure System (LOTS) was first developed. This system was one of the earliest examples in the world of integrated land ownership and valuation records being made available on-line to the whole community, and by the late 1980s over 1000 on-line terminals were available throughout South Australia which gave the public rapid access to detailed property records. For a fee of approximately $5 per screen of information viewed and downloaded, users of the system could obtain information in a matter of seconds that would normally take them days to collect, and it was received from a source that was authoritative and up-to-date, yet located many hundreds of kilometers away. In essence, the usability of the data was due to several factors such as the provision of an entirely **new service**, the **convenience** and **speed of access**, the relatively **low-cost** of the data, its **integrity**, and the fact that even though it was a product derived simultaneously from several government databases its method of **presentation** to the client took the form of a single integrated report. By the early 1990s over 20,000 online queries per day were being received, which in turn were believed to be yielding in excess of $1 million per week to the custodians of the system, indicating a clear level of **satisfaction** with the information products being sold.

Similarly, there is no doubt that datasets which are **authoritative**, due to either the **data producer's reputation** or by being **officially sanctioned** by public agencies, develop a level of usability which enhances their value to consumers. This may not necessarily mean they are guaranteed free of error (which may constitute another element of usability), but official sanction promotes confidence amongst users that they will be using data which are widely adopted in their particular industry, and that its use will permit them to more easily cooperate with others applying the same datasets. The high reputation of a data producer will also produce the same effect, and for many years the British Ordnance Survey maps have been held in high regard worldwide. Sometimes users of spatial information are compelled to use official data sets, for example in defence applications, and it is not so much the authoritative nature of the data that makes them usable but the fact they will conform to **international standards**, and that they are coupled with high **security** measures and their application by allied defence forces will be tightly controlled. On the other hand, it can simply be a case of **popularity** for a particular information product which makes it highly desirable to obtain and use—and the analogy here would be the expression used in the 1980s and 90s that "nobody ever got fired for buying IBM".

Of course, **benefits** cannot be overlooked when discussing information usability and as mentioned earlier these can take social, environmental and economic forms. In the case of Dr John Snow's work in discovering the source of cholera in London, there were obvious life-saving benefits associated with the use of his information, while in the case of discovering the Greenhouse Effect and possible global warming in the 1980s, it was the use of satellite data over the Antarctic that has led to changes in the way we think about our environment. At the same time, coupled with benefits is the issue of **cost** and expensive information would certainly have the potential to impact negatively upon the usability of a dataset. An example of this occurred in Australia in the early 1980s when a water utility sought a fee of over $1million for a copy of its digital land parcel base map which had been requested by a planning authority for use as a backdrop when digitizing its planning zone boundaries. The planning agency simply wanted to use the parcel base to ensure that these two important fundamental datasets would overlay correctly, but the agency baulked at the high cost of the temporary use of the parcel base, and there was some concern (before the matter was eventually resolved) that the agency would opt to digitise its own planning zone map—which would have led to a duplicate (and eventually conflicting) parcel base map being created at unnecessary cost. The issue of cost can also work in reverse where it raises usability, as in the case of highly expensive and **exclusive** business data, where users are willing to pay a premium for information that few others can afford, and often hold the belief that high-cost data has greater **integrity**.

The **type of decision** that will the taken with the data may well impact upon its eventual usability. For example, critical versus non-critical decisions, political versus non-political decisions, and routine versus non-routine decisions, can each require different levels of usability within the data employed for a decision task. And in the case of legally-enforceable decisions, emphasis may be placed less on the cost of the information than its ability to be **legally defensible**. Coupled with this are the issues of **validity** and **reliability**, and an example was given at the start of this paper where it was reported that governments are becoming increasingly hesitant to take high-risk political decisions concerning the environment when scientific information that would support those decisions cannot be trusted.

Clearly, questions of **trust** would seem to have a major impact upon information usability, and the populations of some African countries have such strongly-held beliefs that government-collected household census information will be misused, that they deliberately lie to census takers—with the result that expensive census collection efforts are virtually worthless for public policy making. Trust in a dataset is a highly valued asset for data producers and users alike, and may arise from a variety of actions such as the knowledge that an information product has been developed by a particular organization, official sanction, adherence to standards and **industry best practice**, the description of **product purposes** (e.g. aeronautical and hydrographic datasets), **certification** of datasets that they have been produced to meet the needs of certain users, and through **caveats on** its **content**.

Data quality and **accuracy/freedom from error** would also seem to be essential candidates for inclusion in any list of usability elements since no user wants to apply error-prone data to their tasks. However, spatial data quality can take many forms and different users will invariably have different requirements before their data can be considered acceptable. Apart from **positional** and **attribute accuracy**, **logical consistency** and **completeness**, **temporal accuracy** or **currency** is becoming increasingly important now that the spatial datasets which were so laboriously created in the 1980s and 90s are now being regularly maintained and updated. This also raises the question of **shelf-life**, and spatial data users in some disciplines are beginning to question just how long a dataset can remain up-to-date (or representative of the real-world) before its usefulness expires. Some data producers **guarantee** their information to be correct, which undoubtedly establishes a high level of faith in the information. An example of this can be found in the land parcel databases that record property ownership rights and restrictions in each Australian State and Territory, and which carry a government-backed guarantee that the information contained within them is correct—otherwise compensation is paid for any harm suffered as a result of information error.

Strongly linked to data quality is the subject of **metadata** (or data about the data). Descriptions of datasets have assumed critical importance now that the Internet is being widely used as the most common means of searching for and accessing datasets. However, no-one appreciates receiving large numbers of ‘hits’ to a search query when it is so time-consuming to investigate every one of them, so increasingly metadata is seen as the solution to ensuring that the datasets required by a user are actually **searchable**. Metadata can be expressed at several levels, with discovery metadata being used to identify a potential resource, and exploitation metadata carrying more information about that resource in terms of its nature, including its quality (Craglia & Evmorforpoulou, 2000). An example of poor metadata affecting the usability of a dataset was personally noted by the lead author in the late 1980s, when the option of using an easily available landuse dataset arose. At first glance the landuse categories that were employed seemed ideal for the research purposes, however closer examination of the metadata revealed the fact there was no information relating to how multiple landuse classes in a single land parcel had been dealt with—so there was no choice but to discard the data.

Some applications of spatial data will have an obvious **novelty** that makes them attractive to users, and an example of this occurred in the 1990s with a product called “Which Restaurant?”. It sold for just a few dollars in Australia (and eventually other parts of the world) on a diskette at shopping outlets, and contained a simple form of spatial database that would permit a user to search for restaurants on the basis of attributes such as meal cost, location, type of cuisine, smoking/non-smoking tables, and lunch/dinner availability for each day of the week. A report could be quickly generated and a simple map was easily produced. Novice users rapidly became skilled in using the product without the need for a help manual, and an updated version of the data (which was developed for each major city it was sold in) was available every 6 months. In more recent times, the growth of data mining and knowledge discovery software has meant that users are now examining previously seemingly routine data sets to identify characteristics that reflect newly-found levels of **interestingness** or **unexpectedness**.

Finally, there would seem to be a usability element associated with the capacity of some information systems to allow users, ranging from novices to expert analysts, to easily apply a variety of methods and techniques to their data. **User skill levels** will always vary, and good **software interfaces** and **visualisation tools** are able to take this into account so that users can gain the most information from their data. Whereas the command-line interfaces of the 1980s left no room for error and required detailed knowledge of the software for their operation, the graphic user interfaces of the 1990s increased the usability of many spatial datasets particularly for new users. Perhaps the earliest recorded failure of a system that used spatial data is that of the Domestic Information Display System (DIDS), developed in 1978 as part of the U.S. White House Information System. At the time, it was seen as a major breakthrough in the automatic provision of statistical data in color map form to State and Federal agencies, in which they would go on-line to a centralised statistical database and select data that would be subsequently mapped and downloaded to local output devices. In fact, the technology was revolutionary but as Cowen (1982, pp. 89-90) notes "Five years later DIDS has been essentially mothballed. … DIDS was isolated, expensive, restrictive, inflexible, somewhat tedious and produced maps that were appropriate for only certain functions." Indeed, technically the system was successful but factors that led to its failure are now recognized classic indicators of poor system and data usability.

4 RESEARCH QUESTIONS AND PRIORITIES

What, then, are the research questions and priorities associated with spatial data usability? The obvious starting point is to question whether the elements presented in the previous section are exhaustive, or else simply a beginning in understanding usability. Certainly, the approximately 40 elements listed above would appear to cover the subject, however there could be others such as the ability of effective marketing to make an information product usable.

Second, there is the question of whether or not the elements can be broadly grouped into a distinct set of common classes. It would appear that issues such as data certification, shelf-life, standardisation, quality, trust, reliability and validity each point to a similar goal relating to internal characteristics of the data, whereas elements such as application and decision type, defensibility, and product purpose raise questions about what the data will be used for. Next, there might be a group of elements relating to costs, benefits, service provided, satisfaction, integration and added-value which indicates a commonality relating to how well the data meet the needs of the application. Finally, elements such as software systems, accommodation of varying user skill levels and visualisation tools point towards a common thread associated with ease of application of the data.

Thirdly, we need to question how these various elements impact upon each other and whether there is a hierarchy of usability elements. For instance, can poor data quality rule out the use of a low-cost, easily available dataset which would otherwise appear to be able to produce the desired product? Alternatively, should a top-down approach be used to assess usability, such that application-based elements are considered first, then quality elements and so forth?

Fourth, if these are in fact the elements of usability then what are their metrics? While a range of measures have already been developed to suit information technology usability testing, it must be remembered that most IT systems are designed to achieve a particular task and what the user wants to do with the system is often not open-ended (for example, a bank teller has only a small number of functions available necessary to do her/his job). However, when it comes to using spatial data this situation is quite different, since the intended use of the data usually remains unknown to both the data producer and the software vendor.

Finally, what methods can be used to apply these metrics? Is there an information product equivalent of 'usability evaluation' as applied in computer interface design and testing which seeks to optimize an interface before release to the market?

5 CONCLUSION

This paper has introduced the topic of spatial data usability, and the authors suggest that it is just as appropriate to consider this subject in the context of data and information, as it has been for the past decade in relation to computer system design and testing. Just as information technology is designed to be usable, so should their derived information products, which can be just as costly to develop. To date, very little is known about spatial data usability, but there are now sufficient case histories to suggest there have been both highly successful and unsuccessful information products developed. A preliminary review indicates there are at least 40 elements of usability that relate to spatial information, although these might well be grouped into five distinct similarity classes. Future research questions have been raised relating to what is the full range of usability elements, is there a hierarchy of elements (are some more important than others), what is the effect of conflicting elements upon each other, what are their metrics, and how can they be measured?

6 ACKNOWLEDGEMENTS

The authors acknowledge funding support for this research from Wageningen University and Research Centre (The Netherlands) and the University of Melbourne (Australia).

7 REFERENCES

Alertbox (2000) Mailing List Usability, Jakob Nielsen's Alertbox, August 20, 2000. Retrieved 23 October, 2001, from useit.com: Jacok Nielsen’s website: http://www.useit.com/alertbox/20000820.html.

Alertbox (2001a) Mobile Devices will Soon be Useful, Jakob Nielsen's Alertbox, September 16, 2001. Retrieved 23 October, 2001 from useit.com: Jacok Nielsen’s website: http://www.useit.com/alertbox/20010916.html.

Alertbox (2001b) Success Rate: the Simplest Usability Metric, Jakob Nielsen's Alertbox, February 18, 2001. Retrieved 23 October, 2001 from useit.com: Jacok Nielsen’s website: http://www.useit.com/alertbox/20010218.html.

Bernstein, P.L. (1998) *Against the Gods: The Remarkable Story of Risk*. John Wiley & Sons: New York.

Beven, K. (2000) On Model Uncertainty, Risk and Decision Making. *Hydrological Processes*, 14, 2605-606.

Cowen, D.J. (1982) Rethinking DIDS: the Next Generation of Interactive Color Mapping Systems, *Cartographica*, 21, 89-92.

Craglia, M. and Evmorforpoulou, K. (2000) Potential Benefits of Agreed Metadata for Data Sharing. *Draft Proceedings of the Workshop: Towards Harmonized Information Systems for Mine Action in South Eastern Europe*, (Peckham, R.J. Ed.) European Commission Joint Research Centre, Ispra, Italy, pp. 69-81.

EarthWatch (2001) EarthWatch Weather On Demand. Retrieved November 27, 2001 from the Earthwatch website: http://www.earthwatch.com/.

EPA (2001) 2001 National Site Assessment Conference: Data Usability for Site Assessment Training Syllabus. Retrieved 23 October, 2001 from the US EPA website: http://www.epa.gov/superfund/programs/siteasmt/sa_conf/syllabus.htm.

ESRI (2001) ESRI Usability Engineering: What is Usability? Retrieved 23 October, 2001 from ESRI website: http://www.esri.com/software/usability/whatisusability.html.

Frank, A.U. (2001) How Usable are Current Geoinformation Products? *GIM International, 15*(10), 73.

Government of South Australia (2001) Land and Property Information. Retrieved 4 December, 2001 from the Department for the Environment and Heritage website: http://www.environment.sa.gov.au/mapland/lapi.html.

Hom, J. (1996) The Usability Methods Toolbox. Retrieved 6 November, 2001 from JT Hom’s website: http://jthom.best.vwh.net/usability/usahome.htm.

International Organisation for Standardisation (1998) *ISO 9241-11: 1998 Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability*. Geneva: ISO.

International Organisation for Standardisation (2001) *ISO/IEC 9126-1:2001 Software engineering—Product quality—Part 1: Quality model*. Geneva: ISO.

Jacobsen, N.E. (1999) Usability Evaluation Methods: the Reliability and Usage of Cognitive Walk-through and Usability Tests. *PhD thesis*, University of Copenhagen, Denmark. Retrieved 23 October, 2001 from the Userforum website http://www.userforum.dk/content/afhandlinger/uem.pdf

King, J.H. (1998) White Paper on NASA Science Data Retention. Retrieved 23 October, 2001 from the National Space Science Data Center website: http://nssdc.gsfc.nasa.gov/nssdc/data_retention.html.

Metadata Humane Society (n.d.) Overview of Usability Evaluation Procedures. Retrieved 23 October, 2001, from the Metadata Humane Society's website: http://www.nacse.org/mhs/review/review.procs.html.

NARA (2001) Long-term Usability of Optical Media, U.S. National Archives and Records Administration. Retrieved 23 October 2001 from the Conservation Online website: http://palimpsest.stanford.edu/bytopic/electronic-records/electronic-storage-media/critiss.html.

National Cancer Institute (2001) Usability Basics. Retrieved 23 October, 2001 from the National Cancer Institute website: http://www.usability.gov.

NCGIA (1989a) *Use and Value of Geographic Information: Initiative 4 Specialist Meeting Summary Report, Technical Report 89-6*, National Center for Geographic Information and Analysis (NCGIA), University of California, Santa Barbara.

NCGIA (1989b) *Selective Bibliography: Value of Information, Technical Report 89-8,* National Center for Geographic Information and Analysis (NCGIA), University of California, Santa Barbara.

NCGIA (1993) *Use and Value of Geographic Information: Final Report*, National Center for Geographic Information and Analysis (NCGIA), University of California, Santa Barbara.

NOAA (2001) High-speed Viewing and Access to Digital Orthophotography via the Internet, Retrieved 23 October 2001, from the U.S. National Ocean and Atmospheric Administration (NOAA) website: http://www.hpcc.noaa.gov/props2/990013.htm (no longer available on line).

OMB (2001) Guidelines for ensuring an maximizing the quality, objectivity, utility, and integrity of information disseminated by Federal Agencies. Retrieved 24 October 2001, from the U.S. Office of Management and Budget website: http://www.whitehouse.gov/omb/fedreg/final_information_quality_guidelines.html.

Spool, J. (2001) The Usability of Usability. Retrieved 23 October, 2001 from webword.com website: http://www.webword.com/interviews/spool2.html.

STC (2001) Presidential Election Ballot—Topics in Usability, Retrieved 23 October 2001, from the Usability Special Interest Group, Society for Technical Communication (STC) website: http://www.stcsig.org/usability/topics/ballot/.

System Concepts (2001) The Usability of On-line Banking: How not to do it. Retrieved 8 November 2001, from the systemconcepts website: http://www.system-concepts.com/articles/barclaysonline.html.

UCLA (2001) John Snow: An Historical Giant in Epidemiology, Retrieved 8 November 2001, from the Department of Epidemiology, School of Public Health, University of California website: http://www.ph.ucla.edu/epi/snow.html.

UPA (2001) Homepage of the Usability Professionals' Association (UPA) Available from: http://www.upassoc.org/.

YellowMap AG (2001) Homepage of YellowMap, Available from : http://www.yellowmap.de/.